The Age of Recommendation Systems: Examining Social Risks of Algorithmically Tailored Content

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Kelly Schaefer

Spring 2022

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Advisor

Bryn E. Seabrook, Department of Engineering and Society

STS Research Paper

Considering Users: A Society of Personalization

Relevance. Convenience. Ease. The marketing of technology has been saturated with values of convenience and ease since early twentieth century newspaper advertisements on the labor-saving qualities of household appliances (Fox, 1990). Modern integration of software into virtually all aspects of daily life is accompanied with design goals of seamlessly facilitating user interaction with relevant information. In this context, machine learning systems built to infer user interest from categorical information and collected data are increasingly embedded in modern society's engagement with software and the greater world. When a user intuitively follows the suggestion provided, design goals of convenient access are satisfied. However, it is vital to consider the nuances and impacts of an increasingly ubiquitous dynamic where human-built software mitigates user consumption of media and information. It is vital to question whether "frictionless" is an ethical design value.

In discussions of algorithm design, it is pertinent to consider the power of tailored algorithms to subtly shape user opinions (Benjamin, 2020). Researchers have found that recommendation systems have the propensity to amplify extreme beliefs and machine algorithms reproduce societal bias (Whittaker, Looney, Reed & Votta, 2021; Caliskan, Bryson, Narayanan, 2017). Additionally, filtering through a lens of what is relevant can narrow the diversity of information that a user is exposed to (TED, 2011). These risks are centered around user psychology and decision making. When interacting with a system designed to ease the critical-thinking burden of decisions, do users consider these impacts? Ulrich Beck's proposed method of risk analysis offers an instructive means of assessing the dynamics and impacts of the contemporary recommendation ecosystem (Mythen, 2004). The framework centralizes the importance of user awareness of risk as they interact with personalized content. Additionally, the exploration of risk production underscores a responsibility for technology developers and service providers to offer greater transparency in their use of tailored algorithms. Ethical deployment of technology requires considering the psychosocial impacts of recommendation systems on individual users.

Extra! Extra! Read All About It! Documentary Research of Recommender Risk

Through documentary research methods, the analysis explores social risks identified in recommender systems and how these risks are mitigated by users and institutions (Mogalakwe, 2006). The paper opens by establishing two domains of social risk in the context of recommendation systems: equity risk and polarization risk. The analysis of social risk draws on sources gathered from prior reading in an ethics course, databases for computing research, databases for Science, Technology and Society studies research, databases of sociology research, government reports, news articles, computing codes of ethics, advocacy books and social media policies. After establishing social risks, the paper moves into a discussion on how users mitigate algorithmic harm from recommendation systems. The discussion of user risk mitigation is grounded in research into algorithmic folk theories, research into algorithmic resistance and contemporary evidence. Transitioning into the policy domain, the paper uses the European Union's General Data Protection and Regulation legislation to explore how risk is mitigated at an institutional public policy level.

How Did We Get Here? Background on Recommender Systems

The story of recommendation systems begins at Xerox PARC, the "Silicon Valley research incubator" that brought the world laser printing and Ethernet (Condliffe, 2012). The year was 1992. Fed up with the menial task of sorting through an inundation of work emails, a group of Xerox PARC employees, including Doug Terry, current Vice President of Amazon Web Services, developed Tapestry: “an experimental mail system” (Goldberg, et. al, 1992). This application allowed users to “react” to whether a certain email was interesting or uninteresting. In turn, the system learned which emails to prioritize when filtering a user’s inbox. From Xerox PARC, the method of collaborative filtering was born.

In the next decade, recommendation algorithms saw applications from online discussion boards to music to movies to jokes (Ekstrand, et. al, 2011). By the year 2000, recommender systems had become “an integral part of some e-commerce sites such as Amazon.com and CDNow” (Burke, 2002). In recent years, recommender systems built using machine learning have become “omnipresent in our lives” and the technology has trended towards increased user profiling (Goldenberg et. al, 2021). Academia describes user profiling as “collecting, organizing and inferring … user profile information” where a user profile summarizes “user's interests, characteristics, behaviors, and preferences” (Eke, et. al, 2019). These user profiles inform the decisions made by recommender systems. Recent advances in personalized algorithms utilize deep learning, a branch of machine learning that applies layers of mathematical models to simulate neural networks. Often, the mechanisms of these algorithms are not interpretable by humans. Personalized deep learning is seen in Google’s click-through-rate predictions and in Twitter’s customer purchase predictions (Goldenberg, et al., 2021).

Terms and Conditions: Risk Analysis of Algorithmically Tailored Content

The framework used for sociotechnical analysis of the recommendation ecosystem is drawn from the work of German sociologist Ulrich Beck. In his 1992 conceptualization of a “risk society," Beck discusses growing risks associated with modernity (Beck, 1992). In particular, Beck explores the dynamic between institutional definition of risk and public perception of risk. Central to his conceptualization of modern risk analysis is the friction in risk identification between scientific and social considerations. Institutional definition of risk is characterized as a capitalist "calculus of risk" grounded in actuarial analysis of hazard probability. Social considerations, he worries, are not always represented in these calculations. The intersection of multiple considerations necessitates active public engagement with risk issues. According to Beck, the responsibility for fostering an accurate, informed perception of dangers is placed on institutional actors such as governments, scientists, legal experts and journalists.

Sociologist Gabe Mythen critiques blanket characterizations found in risk analysis (Mythen, 2004). He finds Beck's characterization of public distrust of science to lack sufficient evidence. Additionally, while he concedes that institutions play an instrumental role in the production of risk, he believes the role is more nuanced. The application of risk analysis found in this paper will not focus on distrust of science, but instead will center perceptions of psychosocial risk and the nuanced relationship between big tech corporations and users.

While relevant to discussions on technology, Beck's framework of risk analysis is not unique to the field of STS. The risk society is a broad social theory drawn upon by a variety of academic disciplines from sociology to politics, from criminology to environmental studies (Mythen, 2007). Prior applications of risk analysis provide insight into how to best apply this theory to discourse on personalized recommendation systems.

In a 2011 exploration of protest, environmental scientist Frances Drake applies a risk society framework to an instance of community resistance to mobile phone towers. She asserts that current analysis of resistance to phone towers centers attention on inconsistent identification of health risks. According to Drake, hyper focus on the tension between public perception of health risk and expert ambivalence limits analysis to “concerns legitimated by science” (Drake, 2011). To more fully identify risk, Drake argues, it is necessary to factor in social context. In her essay, Drake leverages scholarship on neoliberalism and risk society to center public voices and contextualize the resistance of rural communities. Drake identifies the threat modernism poses to rural lifestyle as a central risk. Phone towers visually represent the “hectic, modern world”, which disrupts the rural landscape. Drake’s analysis and critiques highlight that human experience of risk is multidimensional. The main points of contention between experts and the public do not fully represent the public’s perception of risk. When analyzing the effects of algorithmic recommendation on users, it is not sufficient to merely examine points of contention between tech corporations and the public, such as issues of transparency. Risk society analysis requires a broad and multidisciplinary scope that examines ontological user experience.

Age of Recommendation: Results and Discussion

Overview of Results

Users and scholars are concerned over the deleterious effects of social stereotypes and biases in the content and distribution patterns of recommender systems. Critics of these content presentation algorithms worry that algorithmic behavior will continue to suppress and marginalize historically minoritized groups. Underrepresentation of marginalized identities in algorithm development and corporate priorities substantiates an under identification of equity risks. Additionally, rising polarization has led governments, news agencies and advocates to identify algorithmic recommendation as a social risk for intellectual isolation from opposing viewpoints. The validity of this concern is debated in academia, as cognitive bias in self-selected content consumption is arguably a greater risk for increased polarization. Regardless, ethical design of recommender algorithms can improve polarization risk. However, social media companies have historically denied responsibility for understanding or mitigating polarization risk. When confronted with risks of algorithmic inequity, users have developed theories of algorithmic function in order to generate strategies for resisting algorithmic harm. Algorithmic resistance strategies have included individual and collective interactions with algorithmic recommendation systems that aim to change algorithm behavior to be more equitable. Users have also sought redress for inequity and polarization risk outside of the recommender environment. Both legal action and advocacy are viewed as avenues to change the design of algorithmic systems. Institutional actors have sought to implement legislation to protect user autonomy in interacting with the risks of recommender algorithms.

Equity Risks: Bias, Stereotypes and Historical Marginalization Oh My!

In the context of recommender systems, reproduction of societal bias by algorithms raises concern around social risk. The societal bias encoded in machine learning algorithms has gained significant attention in recent years (Benjamin, 2020; O'Neil, 2016). In 2017, researchers found bias in semantic associations learned from standard training text (Caliskan, Bryson, Narayanan, 2017). The researchers analyzed the output of a standardly trained machine learning model and found higher degrees of association between White-identifying names and pleasant terms than between Black-identifying names and pleasant terms. In addition, female names had more association with art and family, where male names had more association with science and career

(Caliskan, Bryson, Narayanan, 2017). An underlying mechanism of modern recommender systems is machine learning technologies. When caution is not taken, recommender systems are privy to learning bias implicit in data.

Encoded bias can have discriminatory impact, as exemplified by a 2013 study on Google advertisements. In the study, Dr. Latayna Sweeney investigated public records advertisements after noticing an advertisement for arrest records in the results of a Google search containing her name. In the process of her investigation, she discovered that public records advertisements implying arrest records appeared 25% more often for searches containing Black-identifying names than searches containing White-identifying names, irrespective of the existence of arrest records (Sweeney, 2013). Additional studies into discriminatory impact have found that negative stereotypical representations are reproduced in the suggestions offered by Google's autocomplete feature and Google image search results (Baker & Potts, 2013; Otterbacher, 2017; Makhortykh et al., 2021).

Ulrich Beck's Risk Society identifies contention between the risk considerations of institutions and the social risks identified by individuals. In the case of recommender systems, concerns around social equity and the reproduction of negative bias are socially defined risks identified not only by academia, but also by users of online platforms. A 2020 investigation into user perception of TikTok's algorithm revealed concern around the suppression of marginalized identities in the content promoted on TikTok's "For You" page. The identified axes of marginalization included race, ethnicity, class, ability status, body size, LGBTQ+ identity and social justice affiliation (Karizat, 2021). The developers responsible for creating recommender algorithms are predominantly Caucasian and Asian men, who belong to relatively privileged identity groups (Williams et. al, 2019; US Equal Employment Opportunity Commission, n.d.).

The underrepresentation of other historically marginalized groups in the development process leads to an under-identification of social equity risks (Benjamin, 2020). A focus on productivity metrics, throughput, and novel technologies in software development leads to an industry culture focused on pushing minimal viable products to market with the mantra "fail fast and fail often" (Ordonez & Haddad, 2008; Draper, 2017). This environment of face-paced development does not prioritize the identification of risk nor does it foster reflection on perpetuated historical inequities. Building more equitable algorithms relies on consideration of historical marginalization and its manifestations in the contemporary. Smith-Doerr et al. argue that to see the benefits of diversity in engineering fields, organizations must go beyond representational diversity. Full integration of historically marginalized and underrepresented individuals is paramount. In a fully integrated environment, inequities are recognized; marginalized workers are seen, heard and respected. To achieve full integration, asymmetric power dynamics must be ameliorated (Smith-Doerr et al., 2017).

Polarization Risks in Recommendation: Inspired By Your Browsing History

---

The interaction of recommender systems with user opinion and perspective is identified as another social risk. Politicians and journalists have voiced concern around the power of personalized content to increase polarization and promote extremism (Great Britain, 2019; Council of the European Union, 2020; Roose, 2019; Hawkins et al., 2019). In 2011, Eli Pariser warned that "filter bubbles", a narrowed worldview resulting from personalized content distribution on the internet, would limit user engagement with diverse perspectives (Pariser, 2011). In 2022, 1584 articles on Google Scholar cite Prasier's 2011 book on the concept. While it is clear that polarization is on the rise in the American political landscape, scholarship is mixed on the role of algorithmic personalization (Pew Research Center, 2017).

Recent studies show that news recommender systems do not limit the diversity of information shown to users (Möller et al., 2018). News recommendations are more likely driven by time and date than user behavior (Courtois, Slechten, and Coenen, 2018). These scholars conclude that the risk posed by recommendation algorithms embedded in news websites is lower than popular perception of risk.

However, a series of recent studies exploring Youtube's recommendation algorithm conclude that Youtube's algorithmic behavior promotes extreme content and leads to ideological filter bubbles (O'Callaghan et al., 2015; Cho et al., 2020; Whittaker et al., 2021). In 2015, O'Callaghan et al. conducted a topic-modeling study to explore first degree recommendations on Youtube. They found empirical evidence supporting a hypothesis that Youtube's recommender system created ideological bubbles with English and German language extreme right content. Other studies followed. In 2021, Whiteker et. al conducted an investigation of recommendation systems across Reddit, Gab and Youtube. They concluded that while Reddit and Gab do not promote extreme content, Youtube continues to amplify extremist content.

In 2016, scholars at the University of Amsterdam conclude that empirical evidence warrants concern around the future trajectory of personalization technology (Zuiderveen Borgesius et al., 2016). One cited study indicated that differences in Google search results can shift undecided voters by 20% (Epstein and Robertson, 2015). However, Zuiderveen Borgesius et al. note that contemporary risk is limited by the fact that personalized content is not the sole source of news information for most users. Additionally, the scholars cite projections that ethically developed news personalization can help expose users to more diverse viewpoints than self-selection (Helberger, 2011; Zuiderveen Borgesius et al., 2016).

User content selection is influenced by cognitive bias that affirms existing beliefs. Bakshy, Messing and Adamic argue that influence of modern recommender algorithms is lower than the influence of the user's choices (Bakshy, Messing, & Adamic, 2015). However, Lumb refutes the validity of Bakshy, Messing and Adamic's study (Lumb, 2015). Given knowledge of the detrimental impacts of user's selective behavior patterns, an experiment conducted in 2020 investigated the specific features of the algorithm that lead to intellectual isolation. The researchers drew a conclusion that recommendations based on social factors exposed users to more diverse content than recommendations based on user history (Cho et al., 2020). German scholars proposed a statistical model of a "triple filter bubble" effect to describe the level to which individual, social and technological filtering leads to polarization. This study ties together a number of factors driving polarization risk to map the role of self-selective user psychology, demographic features, and social network composition in algorithmic impact (Geschke et al., 2019).

The contradictory conclusions across scholarship indicate dissonance in risk analysis. Different scholars use different metrics to evaluate whether recommender systems are a risk for filter bubbles and polarization. There is a standard worry that polarization is increasing, but the role of technology in this trend is debated and nuanced. Some scholars are less concerned by the risk posed by algorithmic personalization, due to a greater emphasis on the risk posed by user psychology and self-selective behavior. Though the predominant source of risk production is debated, scholars have argued that developers have a social responsibility to take accountability for the effects of personalized content distribution (Ansgar et al, 2017). Foundational to Ansgar et al.'s call to action is emphasis on algorithmic transparency. According to the ethical code of the Association for Computing Machinery, developers have a responsibility to understand, communicate and manage risk that accompanies their algorithms (Association for Computing Machinery, 2017). Understanding the current effect of algorithms on user belief and behavior is key to developing more ethical systems. However, through terms of service or public denial of responsibility, social media platforms tend to put responsibility for risk management onto users (Ansgar et al, 2017).

Be The Change: User Level Risk Mitigation

For users to mitigate risks associated with recommender algorithms, they must first be aware of these algorithms. In a 2015 study, Eslami et al. explore user perception of algorithmic personalization in the Facebook News Feed. 62.5% of survey respondents did not have an awareness that a recommender system controlled the content seen in the news feed. After getting more familiar with the difference between curated and non-curated content distribution, survey respondents felt more in control and displayed more active engagement with Facebook content (Eslami et al., 2015). To allow users to actively engage with risk assessment, knowledge of algorithmic decisions in content distribution must be available to the public.

Active engagement with algorithms can manifest itself as algorithmic resistance: a term coined to describe user resistance to algorithmic harm through alternative engagement with platforms. Velkouva and Kaun frame algorithmic resistance as user interactions motivated by an intention to repair and correct problematic behavior in algorithms (Velkoua and Kaun, 2021). Examples of algorithmic resistance can be seen in the responses of TikTok users to perceived algorithmic suppression of marginalized identities. Users reported actions such as intentional engagement with content from marginalized creators, collective commenting and sharing of content believed to be suppressed, modification of content creation to exhibit qualities believed to be favored by the TikTok algorithm, and "following sprees" where users intentionally followed accounts of marginalized individuals (Karizat, 2021). Another example of algorithmic resistance is seen in tactics employed by Asian American and Pacific Islander communities to decolonize cultural identity on the Reddit platform (Dosono and Semaan, 2020).

In addition to resistance through digital platforms, users have resisted recommender system algorithmic harm through public advocacy and lawsuits. The group AlgoTransparancy mobilized around the amplification of extremism in Youtube's content recommender algorithm (*AlgoTransparency*, n.d.). In 2019, a different group filed a lawsuit against Youtube for the suppression of LGBTQ+ content in its algorithm (Andrews, 2019). Public pressure has led to a series of public relations moves on the part of social media companies to confront user concerns around algorithmic harm (Williams, 2022; Goodrow, 2021; The Youtube Team, 2020).

In response to the discrepancy in risk identification between technology users and technology companies, users mitigate risks of recommender systems through a variety of different avenues. These responses include algorithmic resistance, public advocacy and legal action.

The GDPR: Policy Level Risk Mitigation

The European Union's General Data Protection and Regulation (GDPR) legislation confronts a variety of risks identified in modern internet services. One such concern is transparency in the use of algorithmic decision making. When users are not aware of the use of algorithms, they do not have autonomy to make their own risk decisions. Article 5 of GDPR requires that user's personal data must be "processed lawfully, fairly and in a transparent manner in relation to the data subject" (GDPR, 2018). Whether the data used in algorithmic personalization qualifies as personal data has been the subject of legal debate. According to Article 4 of the GDPR, personal data is defined as data related to an identified or identifiable individual. Lawyers have argued that algorithmic inferences are personal data, thus recommender algorithms have mandated transparency. However, European Courts of Justice have come to conflicting decisions on this classification (Mann & Matzner, 2019).

Another section of GDPR relevant to recommender systems is Article 22. This section mandates that citizens have the right to opt out of decisions made by automated processing (GDPR, 2018). While Article 22 seems to mandate consent in the use of recommendation, companies and lawyers have argued a loophole. The legal teams assert that article does not apply, since decisions for content distribution are only partially based on automated processing.

Risk identification by political institutions does not directly correlate to protection from algorithmic harm. Even when protections exist in policy, implementation by technology companies can lag behind. In a 2019 study, Vedder & Dewitte explored the extent to which it is possible to exercise GDPR rights (Vedder & Dewitte, 2019).

It is also important to note that legal GDPR protections do not extend to US citizens. However, the availability of GDPR compliant technologies in US marketplaces allow users more control over their adoption of algorithmic risk.

Not the Whole Picture: Research Limitations

User identified risk in the context of recommendation systems is a broad field. It was not within the scope of the research timeline to fully explore all domains of risk posed by recommender systems. Instead, the analysis focused on a comprehensive exploration of two domains of risk production. Furthermore, while the research provided an overview of risk identification and mitigation, specific statistics on the prevalence of user concern around each algorithmic risk would benefit analysis.

To Be Continued: Future Research Directions

A domain of risk production left largely untouched in the research was the interaction of recommendation systems with user identity. Concern that recommendation systems reinforce negative aspects of a user's identity, such as racist beliefs or eating disorder tendencies, are articulated in scholarship (Karizat, 2021). Additionally, consumption of media is integral to how individual's form their self-concept. Representations on social media have been cited as a source of ethno-racial identity development (Mims, 2019). Future research to investigate how identity risks manifest in algorithmic interactions would further knowledge on algorithmic environments. Understanding how users can best mitigate identity risk benefits both users and ethical developers.

Exploring the efficacy of different risk mitigation strategies would benefit users in their engagement with algorithms. An analysis of which user and policy decisions are most successful at affecting algorithmic change is fruitful for the future of risk mitigation. Finally, an ethnographic study of the process of algorithmic risk identification in technology companies would provide insight into bridging the discrepancy in risk definition between technology companies and users.

Drawing It All Together: A Conclusion

Underrepresentation of marginalized identities in technology companies has led to development processes and corporate cultures that do not sufficiently identify social equity risks of recommender behavior. Users and academia have raised concerns around encoded bias and the propensity of algorithmic recommendation systems to generate harm by suppressing marginalized voices and furthering negative stereotypes. Users have confronted these risks through advocacy, legal action, and strategic interactions with algorithms aimed to repair algorithmic harm. Additionally, growing concern around the role of algorithmic systems in creating filter bubbles, where users are only exposed to reinforcing viewpoints, has generated a significant level of discourse in academia around polarization risk. Self-selective user behavior is shown to create polarized climates and there is worry that some recommender algorithms amplify this trend. While codes of ethics call for accountability in regards to algorithmic impact, social media companies have historically denied responsibility for the effects of their algorithms. Policy measures, designed to support user autonomy and mitigate algorithmic risk through transparency, have met resistance from corporate legal teams that argue loopholes. Denial of risk responsibility by technology companies places the burden of risk identification and mitigation on users. While users and government institutions have developed strategies to mitigate risk, the under prioritization of risk mitigation by technology companies creates an environment where it is imperative for users to critically consider the social effects of technological systems.

Works Cited:

AlgoTransparency. (n.d.). AlgoTransparency. Retrieved March 28, 2022, from https://algotransparency.org/

Andrews, J. (2019, August 14). *LGBTQ YouTubers Are Suing YouTube over Alleged Discrimination*. The Verge. https://www.theverge.com/2019/8/14/20805283/lgbtq-youtuber-lawsuit-discrimination-alleged-video-recommendations-demonetization

Ansgar, K., Perez, V. E., Helena, W., Menisha, P., Sofia, C., Marina, J., & Derek, M. (2017). Editorial responsibilities arising from personalization algorithms. *The ORBIT Journal*, *1*(1), 1–12. doi:10.29297/orbit.v1i1.26

Association for Computing Machinery. 2017. "Principles for Algorithmic Transparency and Accountability." https://www.acm.org/binaries/content/assets/pub lic-policy/2017_joint_statement_algorithms.pdf

Baker, P., & Potts, A. (2013). 'Why Do White People Have Thin Lips?' Google and the Perpetuation of Stereotypes Via Auto-Complete Search Forms. *Critical Discourse Studies*, *10*(2), 187–204. https://doi.org/10.1080/17405904.2012.744320

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, *348*(6239), 1130–1132. doi:10.1126/science.aaa1160

Benjamin, R. (2020). Race After Technology: Abolitionist Tools for the New Jim Code. Polity.

Burke, R. (2002). Hybrid Recommender Systems: Survey and Experiments. *User Modeling and User-Adapted Interaction*, *12*(4), 331–370. doi:10.1023/A:1021240730564

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, *356*(6334), 183–186. doi:10.1126/science.aal4230

Cho, J., Ahmed, S., Hilbert, M., Liu, B., & Luu, J. (2020). Do Search Algorithms Endanger Democracy? An Experimental Investigation of Algorithm Effects on Political Polarization. *Journal of Broadcasting & Electronic Media*, *64*(2), 150–172. https://doi.org/10.1080/08838151.2020.1757365

Courtois, C., Slechten, L., & Coenen, L. (20181001). Challenging Google Search filter bubbles in social and political information: Disconforming evidence from a digital methods case study. *Telematics And Informatics*, *35*(7), 2006–2015. https://doi.org/10.1016/j.tele.2018.07.004

Council of the European Union. (2020). The Role of Algorithmic Amplification in Promoting Violent and Extremist Content and its Dissemination on Platforms and Social Media.

Dosono, B., Semaan, B.: Decolonizing tactics as collective resilience: identity work of aapi communities on reddit. Proceed. ACM Human Comput. Interact. **4**(CSCW1), 1–20 (2020)

Drake, F. (2011). Protesting Mobile Phone Masts: Risk, Neoliberalism, and Governmentality. Science, Technology, & Human Values, 36(4), 522–548. https://doi.org/10.1177/0162243910366149

Draper, N. (2017). Fail Fast: The Value of Studying Unsuccessful Technology Companies. *Media Industries Journal*, *4*(1). https://doi.org/10.3998/mij.15031809.0004.101

Ekstrand, M. D., Riedl, J. T., & Konstan, J. A. (2011). *Collaborative filtering recommender systems*. Now Publishers Inc.

Eslami, M., Rickman, A., Vaccaro, K., Aleyasen, A., Vuong, A., Karahalios, K., Hamilton, K., & Sandvig, C. (2015). "I always assumed that I wasn't really that close to [her]": 33rd Annual CHI Conference on Human Factors in Computing Systems, CHI 2015. *CHI 2015 - Proceedings of the 33rd Annual CHI Conference on Human Factors in Computing Systems*, 153–162. https://doi.org/10.1145/2702123.2702556

Fox, B. J. (1990). Selling The Mechanized Household:: 70 Years of Ads in Ladies Home Journal. Gender & Society, 4(1), 25–40. https://doi.org/10.1177/089124390004001003

General Data Protection Regulation (GDPR). 2018. *General Data Protection Regulation (GDPR) – Final text neatly arranged*. [online] Available at: <https://gdpr-info.eu/> [Accessed 20 March 2022].

Geschke, D., Lorenz, J., & Holtz, P. (2019). The triple-filter bubble: Using agent-based modelling to test a meta-theoretical framework for the emergence of filter bubbles and echo chambers. *British Journal of Social Psychology*, *58*(1), 129–149. https://doi.org/10.1111/bjso.12286

Goodrow, C. (2021, September 15). *On YouTube's recommendation system*. Blog.Youtube. https://blog.youtube/inside-youtube/on-youtubes-recommendation-system/

Goldberg, Nichols, D., Oki, B. M., & Terry, D. (1992). Using collaborative filtering to weave an information tapestry. Communications of the ACM, 35(12), 61–70. https://doi.org/10.1145/138859.138867

Goldenberg, D., Kofman, K., Albert, J., Mizrachi, S., Horowitz, A., & Teinemaa, I. (2021). Personalization in Practice: Methods and Applications. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining* (bll 1123–1126). doi:10.1145/3437963.3441657

Great Britain, Department for Digital, C., Media and Sport, Great Britain, & Home Office. (2019). *Online Harms White Paper*.

Hawkins, S., Yudkin, D., Juan-Torres, M., & Dixon, T. (2019). *Hidden Tribes: A Study of America's Polarized Landscape* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/xz25v

Karizat, N., Delmonaco, D., Eslami, M., & Andalibi, N. (2021). Algorithmic Folk Theories and Identity: How TikTok Users Co-Produce Knowledge of Identity and Engage in Algorithmic Resistance. *Proc. ACM Hum. -Comput. Interact.*, *5*(CSCW2). doi:10.1145/3476046

Lumb, D. (2015). Why scientists are upset about the Facebook Filter Bubble story. Retrieved from: http://www.fastcompany.com/3046111/fast-feed/why-scientists-are-upset-over-the-facebook-filter-bubble-study

Mann, M., & Matzner, T. (2019). Challenging algorithmic profiling: The limits of data protection and anti-discrimination in responding to emergent discrimination. *Big Data & Society*, *6*(2), 2053951719895805. https://doi.org/10.1177/2053951719895805

Makhortykh, M., Urman, A., & Ulloa, R. (2021). Detecting Race and Gender Bias in Visual Representation of AI on Web Search Engines. In L. Boratto, S. Faralli, M. Marras, & G.

Stilo (Eds.), *Advances in Bias and Fairness in Information Retrieval* (pp. 36–50). Springer International Publishing. https://doi.org/10.1007/978-3-030-78818-6_5

Mims, L. C. (2019). “Our Year to Shine:” Exploring How Media Affects Black Girls. *Conference Papers -- American Sociological Association*, 1–25.

Möller, J., Trilling, D., Helberger, N., & van Es, B. (2018). Do not blame it on the algorithm: an empirical assessment of multiple recommender systems and their impact on content diversity. *Information, Communication & Society*, *21*(7), 959–977. doi:10.1080/1369118X.2018.1444076

Monti, D., Rizzo, G., & Morisio, M. (2021). A systematic literature review of multicriteria recommender systems. *Artificial Intelligence Review*, *54*(1), 427–468. doi:10.1007/s10462-020-09851-4

Mythen, G. (2004). Defining Risk. Ulrich Beck: A Critical Introduction to the Risk Society. (pp. 53-73). London, England. Sterling, Virginia. Pluto Press.

Mythen, G. (2007). Reappraising the Risk Society Thesis: Telescopic Sight or Myopic Vision? Current Sociology, 55(6), 793–813. https://doi.org/10.1177/0011392107081986

O’Callaghan, D., Greene, D., Conway, M., Carthy, J., & Cunningham, P. (2015). Down the (White) Rabbit Hole: The Extreme Right and Online Recommender Systems. Social Science Computer Review, 33(4), 459–478. https://doi.org/10.1177/0894439314555329

O'Neil, C. (2016). *Weapons of Math destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown.

Ordonez, M. J., & Haddad, H. M. (2008). The State of Metrics in Software Industry. *Fifth International Conference on Information Technology: New Generations (Itng 2008)*, 453–458. https://doi.org/10.1109/ITNG.2008.106

Otterbacher, J., Bates, J., & Clough, P. (2017). Competent Men and Warm Women: Gender Stereotypes and Backlash in Image Search Results. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 6620–6631. Presented at the Denver, Colorado, USA. doi:10.1145/3025453.3025727

Pariser, E. (2011). *The filter bubble: How the new personalized web is changing what we read and how we think*. Penguin.

Pew Research Center. (2017, October 5). The Partisan Divide on Political Values Grows Even Wider [Report]. *Pew Research Center - U.S. Politics & Policy*. https://www.pewresearch.org/politics/2017/10/05/the-partisan-divide-on-political-values-grows-even-wider/

Rader, E., & Gray, R. (2015). Understanding User Beliefs About Algorithmic Curation in the Facebook News Feed. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 173–182. https://doi.org/10.1145/2702123.2702174

Roose, K. (2019, June 8). The Making of a YouTube Radical. *The New York Times*. https://www.nytimes.com/interactive/2019/06/08/technology/youtube-radical.html

Smith-Doerr, L., Alegria, S. N., & Sacco, T. (2017). How Diversity Matters in the US Science and Engineering Workforce: A Critical Review Considering Integration in Teams, Fields, and Organizational Contexts. *Engaging Science, Technology, and Society*, *3*, 139–153. https://doi.org/10.17351/ests2017.142

Sweeney, L. (2013). Discrimination in online ad delivery. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2208240

TED. (2011). Tedtalks: Eli Pariser--Beware Online Filter Bubbles. Retrieved October 24, 2021, from https://www.youtube.com/watch?v=B8ofWFx525s.

The Youtube Team. (2020, October 15). *Managing harmful conspiracy theories on YouTube*. Blog.Youtube. https://blog.youtube/news-and-events/harmful-conspiracy-theories-youtube/

U.S. Equal Employment Opportunity Commission, Diversity in High Tech (n.d.). Washington, DC.

Vedder, A., & Dewitte, P. (2019). *Exercising GDPR data subjects' rights*. 100.

Velkova, J., & Kaun, A. (2021). Algorithmic resistance: Media practices and the politics of repair. *Information, Communication & Society*, *24*(4), 523–540. https://doi.org/10.1080/1369118X.2019.1657162

Whittaker, J., Looney, S., Reed, A., & Votta, F. (2021). Recommender systems and the amplification of extremist content. *Internet Policy Review*, *10*(2). https://policyreview.info/articles/analysis/recommender-systems-and-amplification-extremist-content

Williams, J. D., Lopez, D., Shafto, P., & Lee, K. (2019). Technological Workforce and Its Impact on Algorithmic Justice in Politics. Customer Needs and Solutions, 6(3), 84–91. doi:10.1007/s40547-019-00103-3

Williams, M. (2022, February 24). Building the Metaverse With Diversity and Inclusion from the Start. *Meta*. https://about.fb.com/news/2022/02/building-the-metaverse-with-diversity-and-inclusion-from-the-start/

Zuiderveen Borgesius, F. J., Trilling, D., Möller, J., Bodó, B., de Vreese, C. H., & Helberger, N. (2016). Should we worry about filter bubbles? *Internet Policy Review*, *5*(1). https://doi.org/10.14763/2016.1.401